Available online at www.sciencedirect.com
ScienceDirect
Computers and Mathematics with Applications 55 (2008) 1044–1053
An International Journal
computers & mathematics
with applications
www.elsevier.com/locate/camwa

Multiple folding and packing in DNA modeling

Renzo L. Ricca[a,b,*], Francesca Maggioni[a]

[a] *Department of Mathematics and Applications, University of Milano-Bicocca, Via Cozzi 53, 20125 Milano, Italy*
[b] *Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK*

Abstract

In this paper we introduce and analyze a set of equations to study geometric and energetic aspects associated with the kinematics of multiple folding and coiling of closed filaments for DNA modeling. By these equations we demonstrate that a high degree of coiling may be achieved at relatively low energy costs through appropriate writhe and twist distribution, and independently from the number of coils formed. For sufficiently high twist we show that coiling is actually favoured by elastic energy relaxation, when the deformation energy is due to curvature and the mean twist of the filament. We give estimates for the writhing process based purely on curvature information and the number of coils produced. We also determine the packing rate associated with filament compaction in the case of a hierarchical helical coiling. These results find useful applications in DNA biology, especially in modeling DNA wrapping in proteic regions, where there is a strong connection between high coiling, efficient compaction and energy localization.

Keywords: DNA writhing; Coiling; Deformation energy; Packing

1. DNA folding and packing in nature

In this paper we introduce a set of equations to study geometric and energetic aspects associated with the kinematics of multiple folding and coiling of closed filaments for DNA modeling. These equations prescribe the evolution of an inextensible thin filament that evolves in space from a plane circle to a coiled configuration. Folding and coiling are measured by global geometric quantities such as the writhing number and normalized total curvature, while the energetics is based on deformation energy due to the thin rod approximation of linear elastic theory, under conservation of linking number. By using these equations we demonstrate that a high degree of coiling may be achieved at relatively low energy costs through appropriate folding and twist distribution, and independently from the number of coils formed. We determine the deformation energy associated with the coiling process and show that, above certain critical twist values, coiling is actually favoured by the relaxation of elastic energy. We also give estimates for filament writhing purely on the basis of curvature information and the number of coils, and evaluate the packing rate associated with a hierarchical helical coiling of the filament in space. These results shed new light on the folding mechanisms and associated energy contents and, as we shall discuss below, find useful applications in the general context of structural genomics and proteomics.

* Corresponding author at: Department of Mathematics and Applications, University of Milano-Bicocca, Via Cozzi 53, 20125 Milano, Italy.
E-mail address: renzo.ricca@unimib.it (R.L. Ricca).

0898-1221/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2006.12.084

Geometric and topological aspects play a fundamental and yet mysterious role in DNA biology. That morphological and structural properties influence biological functions is an established fact, but a clear understanding of the relationship between shape and function is still a big challenge for the future. In recent years, however, the perception that at various levels of investigation simple geometric and topological modeling helps to detect factors that may be relevant for functional genomics and proteomics has grown considerably. As recent progress in the experimental and computational molecular biology of protein folding highlights [1], simple geometric modeling based on simplified representations of the polypeptide chain are, to some extent, more likely to be successful in predicting both protein structures and protein-folding mechanisms than detailed analysis of the amino-acid sequence. Independent lines of investigation indicate that, indeed, protein-folding rates and mechanisms are largely determined by the protein's morphology rather than by its inter-atomic interactions, as confirmed by the latest results based on statistical coupling analysis [2].

The interplay of structural genomics and proteomics is also evidenced by geometry: small adjustments in bending and twisting of the DNA neighboring base-pairs has an effect higher up on the chemical sequence, both locally, since regions of rapid change in writhing density may identify protein binding sites, and globally, on the overall features of DNA folding. In turn, DNA bending and wrapping (around RNA polymerases in initiation mechanisms, for instance) appear to be a major and general strategy to help bend DNA through the active site and to untwist the helix by the action of transcriptional regulators [3]. Thus, the binding of proteins and other ligands induces a strong deformation of the DNA structure. Since regions susceptible to protein binding or enzymatic attack are regions of concentrated deformation, our ability to locate specific geometric properties could flag physiological activities. Computer simulations based on Monte Carlo analysis and elastic thin rod models (see, for example, [4]), and classical geometric analysis (see, for example, [5]) address precisely these issues.

Geometric information also plays a crucial role in relation to topology. Since the original, systematic studies of knots formed in nicked DNA rings by *Escherichia coli* DNA topoisomerases of type 1 and 2 [6–8], much has been learnt about the changes in molecular geometry and topology introduced by these regulatory enzymes. Topoisomerase of type 1, for instance, acts at intersections of DNA segments (termed nodes) via reversible, transient single-strand breaks by passing either a duplex or single-stranded segment through an enzyme-bridged single-strand break, hence inducing a step-of-one change in the linking number Lk, which measures the number of times that the two DNA strands are intertwined. Topoisomerase 2, on the other hand, acts via double-strand breaks by cutting and rejoining two sugar-phosphate backbones in a process called sign inversion, reducing in steps of two the linking number Lk. The reduction in Lk is then manifested through a change in the overall folding of the double helix and/or by a change in the local twisting of base-pairs. These enzymes, by carrying out DNA strand passage reactions, fulfill a vital role in the replication transcription and recombination of double helical DNA (see the collection of papers in [9]).

Another crucial aspect regards DNA compaction and relationships between energy localization and structural complexity [10]. Enzyme activated strand passages may free-up or lock mechanical energy on site, stimulating or inhibiting further folds and twists of the double helix. Geometric and topological adjustments, brought in by the new changes in curvature, twist and entanglement, determine a change in deformation energy that has profound effect on the mobility of DNA in the cell. If supercoiling is Nature's smart way to compactify a lot of information in a small volume, enhancing compaction through entanglement may induce faster mobility by reducing considerably the radius of gyration of the polymer system [11]. Geometric modeling of folding processes allows us to test and understand possible mechanisms of compaction and estimate packing rates for energy considerations.

This paper is organized as follows. In the next section we introduce a kinematic model for multiple folding and coiling of a closed curve in space and discuss geometric measures of folding and coiling; we show how high coiling may be achieved at low writhing contents and give estimate of writhe in terms of curvature information and the number of coils produced. In Section 3 we calculate the deformation energy associated with coiling, by using linear elastic theory and thin rod approximation, and we show that, for sufficiently high linking number due to the initial twist, elastic relaxation actually favours coiling. In Section 4 we give a simple estimate of the packing rate associated with hierarchical helical coiling of the filament in space, and we conclude with some comments.

2. Modeling and measuring filament folding

A simple, but straightforward, analysis of filament folding is performed by kinematic models of the time evolution of an inextensible, closed curve $\mathcal{C}$ in three-dimensional space. This curve, thought of as the central axis of a

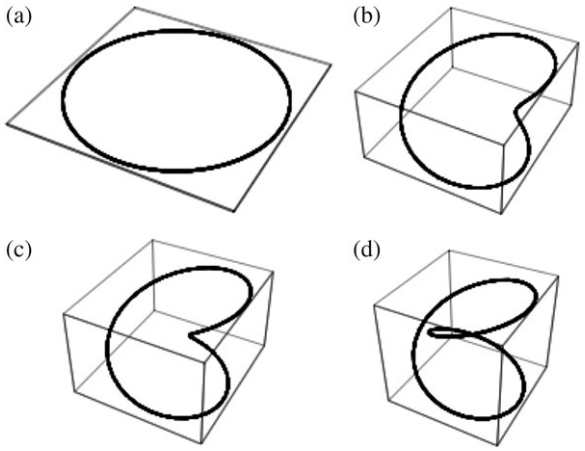

Fig. 1. Single coil formation generated by Eqs. (2) for $m = 1$ and $n = 2$. (a) The initial circular configuration ($t = 0$) (b) writhes up and (c) folds in space, (d) to produce one final coil ($t = 6$).

closed, double-strand DNA filament (as for catenanes; [8]), is given by a smooth (at least C^3), simple (i.e. non-self-intersecting), closed curve $\mathbf{X} = \mathbf{X}(\xi)$ in $\mathbb{R}^3$, where $\xi \in [0, 2\pi]$ parametrizes the points on the curve along its length L. That is

$$\mathcal{C} = \mathrm{Im}(\mathbf{X}), \qquad \mathbf{X} : [0, L] \to \mathbb{R}^3, \tag{1}$$

with $\mathbf{X}(0) = \mathbf{X}(L)$. We take $\mathcal{C}$ to be inextensible, that is, $L =$ constant. The geometry of the axis is prescribed by curvature $c = c(\xi)$ and torsion $\tau = \tau(\xi)$ of $\mathbf{X}$ through the standard Frenet–Serret formulae.

Different folding mechanisms may be explored by a family of time-dependent curves $\mathbf{X} = \mathbf{X}(\xi, t)$ given, for example, by combining a base curve $\mathbf{Y} = \mathbf{Y}(\xi, t)$ with a second curve $\mathbf{Z} = \mathbf{Z}(\xi, t)$ responsible for coiling, so as to have $\mathbf{X} = \mathbf{Y} + \mathbf{Z}$. Here the base curve may represent the secondary structure of the macromolecule, whereas $\mathbf{Z}(\xi, t)$ mimics the tertiary structure. This approach can be extended to include higher-order deformations.

Let us consider the following set of governing equations:

$$\mathbf{X} = \mathbf{X}(\xi, t) : \quad \begin{cases} x = [\cos(m\xi) - t\cos(n\xi)]/l(t) \\ y = [\sin(m\xi) - t\sin(n\xi)]/l(t) \\ z = t\sin(\xi)/l(t), \end{cases} \tag{2}$$

where, in order to ensure inextensibility, each $\mathbf{X}$-component is normalized by the length function

$$l(t) = \frac{1}{2\pi}\int_0^{2\pi} \left[\left(\frac{\partial x}{\partial \xi}\right)^2 + \left(\frac{\partial y}{\partial \xi}\right)^2 + \left(\frac{\partial z}{\partial \xi}\right)^2\right]^{1/2} \mathrm{d}\xi. \tag{3}$$

This re-scaling ensures that the total length is kept fixed at $L = L(0) = 2\pi$.

For simplicity, we choose the base curve $\mathbf{Y} = \mathbf{Y}(\xi)$ in Eqs. (2) to be a plane circle fixed with time, and coiling generated by a deformation given by a time-dependent prescription of the second curve $\mathbf{Z}(\xi, t)$. The parameters $n > m > 0$ are integers that control the total number of coils produced. For $n = 2$ and $m = 1$ we have the formation of a single coil, as shown in Fig. 1. As we see, the coil is produced by folding (writhing) the curve in space, through the development of a loop region that gives rise to the full coil. This type of deformation is known as a Reidemeister type I move, leaving the curve topology unchanged.

An example of multiple folding is shown in Fig. 2, where now seven coils are produced simultaneously. By varying m and n, we have higher degree of complexity: for $m = n - 1$, for example, the first $m - 1$ coils are instantly produced from m coverings of the plane circle, while the remaining coil is generated during the full evolution of the curve. A detailed analysis of the general set of equations, which includes Eqs. (2) as a particular case, is carried out in the companion paper by Maggioni and Ricca [12]. Here we would just like to point out that a suitable choice of the base curve allows us to model rather complex morphologies, like the localized wrapping of DNA around proteic regions

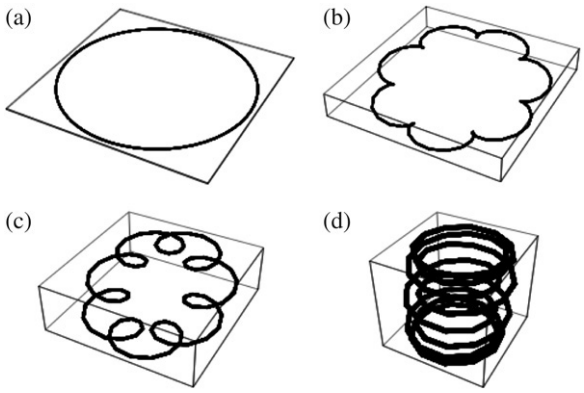

Fig. 2. Formation of 7 coils generated by Eqs. (2) for $m = 1$ and $n = 8$. (a) The initial circular configuration (b) writhes up and (c) folds simultaneously in 7 distinct regions, evenly distributed along the curve, (d) to produce 7 final coils. Note that since the curve is inextensible, folding and coiling determine compaction in the ambient space.

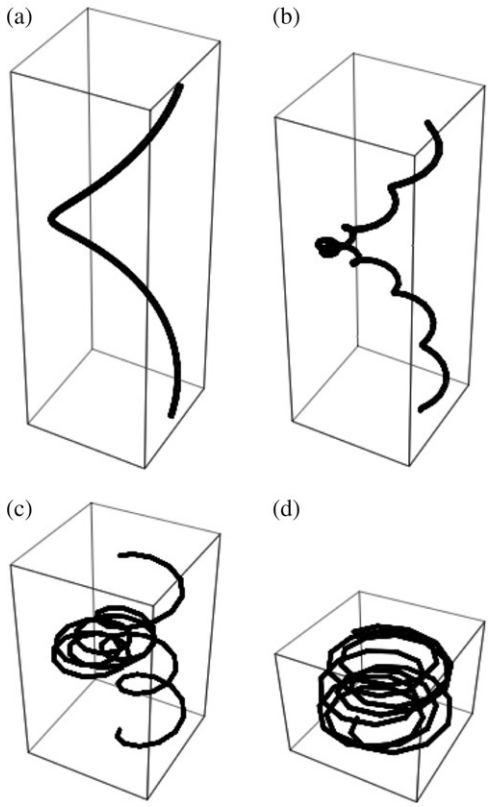

Fig. 3. Formation of seven coils generated on a helical base curve for $m = 1$ and $n = 8$. (a) The initial circular helix ($t = 0$) first (b) writhes up and then (c) folds locally, to form (d) seven coils. Because of the constraint on total length, multiple coiling produces a visible contraction of the curve in space.

by helical multiple folding; Fig. 3 shows a simple example given by taking the base curve to be the standard helix $\mathbf{Y} = (\cos\xi, \sin\xi, \xi)$. More accurate modeling of multiple folding and coiling in complex structures can thus be prescribed *ad hoc* by an appropriate choice of $\mathbf{Y}$ and $\mathbf{Z}$.

Folding and coiling of filaments are measured by the writhing number Wr and the normalized total curvature K, respectively. The writhing number Wr [13] is defined by

$$Wr \equiv \frac{1}{4\pi} \oint_C \oint_C \frac{\hat{\mathbf{t}}(\xi) \times \hat{\mathbf{t}}(\xi^*) \cdot [\mathbf{X}(\xi) - \mathbf{X}(\xi^*)]}{|\mathbf{X}(\xi) - \mathbf{X}(\xi^*)|^3} \|\mathbf{X}'(\xi)\| \|\mathbf{X}'(\xi^*)\| \mathrm{d}\xi \mathrm{d}\xi^*, \tag{4}$$

where $\mathbf{X}(\xi)$ and $\mathbf{X}(\xi^*)$ denote two points on the axis for any pair $\{\xi, \xi^*\} \in [0, 2\pi]$, and $\hat{\mathbf{t}}(\xi) \equiv \mathbf{X}'(\xi)/\|\mathbf{X}'(\xi)\|$ is the unit tangent to $\mathcal{C}$ at ξ. The writhing number admits a physical interpretation in terms of the average number of apparent crossings made by the filament strands in space [13,14]. Suppose that we view the curve $\mathcal{C}$ along a viewing direction $\boldsymbol{\nu}$: let $n_+(\boldsymbol{\nu})$ be the number of positive crossings and $n_-(\boldsymbol{\nu})$ the number of negative crossings seen along $\boldsymbol{\nu}$; then, we have

$$Wr = \langle n_+(\boldsymbol{\nu}) - n_-(\boldsymbol{\nu}) \rangle, \tag{5}$$

where the angular brackets denote averaging over all directions $\boldsymbol{\nu}$ of projection. For example, when the filament writhes up and folds to one coil, as in Fig. 1, a loop region emerges, exhibiting an apparent crossing (Fig. 1(d)). To each crossing we can assign a positive or negative sign according to the orientation of the filament strands induced by the tangent to $\mathcal{C}$. The crossing may be more or less visible, according to the viewing direction and the minimal distance between the strands; hence, Wr is not necessarily an integer, varying continuously with the continuous deformation of the curve in space. By applying the average crossing number interpretation to the case of Fig. 1, we can see that Wr must necessarily grow from 0 (for the plane circle) to the asymptotic value of 1, attained when the coil is fully formed as $t \to \infty$.

Coiling can be measured by the normalized total curvature of $\mathcal{C}$, which is given by

$$K \equiv \frac{1}{2\pi} \oint_{\mathcal{C}} c(\xi) \|\mathbf{X}'(\xi)\| \mathrm{d}\xi \ . \tag{6}$$

For the plane circle of radius $R = 1$, coiling has a lower bound given by $K = 1$.

The folding process conserves topology. In the case of a thin filament, this means conservation of the linking number Lk, according to the well-known formula [15,16]

$$Lk = Wr + Tw, \tag{7}$$

where Tw represents the total twist of the filament fibers along $\mathcal{C}$. Denoting by $\Omega = \Omega(\xi)$ the angular twist rate of the fibers, we have

$$Tw \equiv \frac{1}{2\pi} \oint_{\mathcal{C}} \Omega(\xi) \|\mathbf{X}'(\xi)\| \mathrm{d}\xi \ , \tag{8}$$

that is related to the geometry of the filament axis through the decomposition (see, for example, [14])

$$Tw = \frac{1}{2\pi} \oint_{\mathcal{C}} \tau(\xi) \|\mathbf{X}'(\xi)\| \mathrm{d}\xi + \frac{1}{2\pi} [\Theta]_{\mathcal{F}} = \mathcal{T} + \mathcal{N} \ , \tag{9}$$

where the first term in the r.h.s. of (9) is the normalized total torsion $\mathcal{T}$ and the second term is the normalized intrinsic twist $\mathcal{N}$ of the filament fibers *around* $\mathcal{C}$. For simplicity we take the fibers to be closed curves, wound uniformly around $\mathcal{C}$; thus $\mathcal{N}$ is an integer.

Since during deformation Lk remains constant, a change in writhing number Wr must be compensated by an equal and opposite change in the total twist Tw. Hence, conservation of linking number imposes a strong constraint on possible writhing rates, since these are ultimately related to the rates of total twist that are admissible by a physical system (see Section 3 below). However, high coiling may be also achieved at relatively low writhing values; as mentioned earlier, since Wr is related to the average number of *signed* crossings, the generation of multiple coils of alternating, opposite sign implies a bound on the writhing number to low values. Eqs. (2) have actually the property to generate coils of alternating sign: for $m = 1$, and for any even number n, we indeed have the production of $N = n - 1$ coils of alternating, opposite sign, whose contributions, on average, cancel out, maintaining the writhing number bounded between 0 and 1 (as in Fig. 4(c)), and when $m = n - 1$ with $m > 1$ integer, then $Wr \in [0, m]$. The

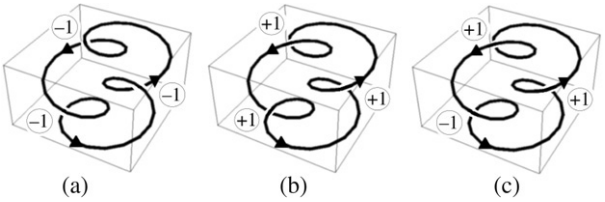

Fig. 4. Estimates of writhing number based on the average crossing number interpretation, for three different cases of three-coil formation: (a) $Wr \approx -3$; (b) $Wr \approx +3$; (c) $Wr \approx +1$.

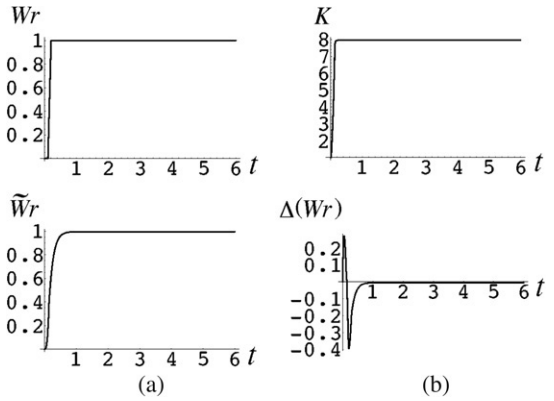

Fig. 5. Top row: comparison between (a) writhing number $Wr = Wr(t)$ and (b) normalized total curvature $K = K(t)$ for the seven coils produced by Eqs. (2), plotted against time. Bottom row: (a) $\widetilde{W}r = h_n(t)K(t)$ and (b) absolute error $\Delta(Wr) = Wr(t) - \widetilde{W}r(t)$, for the seven-coil case, plotted against time. The localized jump in $\Delta(Wr)$ is due to the instantaneous formation of the coils.

conservation of linking number, then, implies (by Eq. (7)) that a small change in the writhing number corresponds to a small change in total twist, which is crucially important for energy considerations (see Section 3 below).

The total coiling induced by filament folding is measured by the growth in K, due to the average increase in curvature associated with the formation of loops. Hence, both Wr and K increase with time. A comparison between the two growths for varying n (with $m = 1$) shows a remarkable similarity, especially in the case of high coiling, as evidenced by the top diagrams of Fig. 5. The difference in absolute value is due to the fact that, while $Wr \in [0, 1]$, K, which in principle can grow indefinitely, is actually limited by the number of coils formed, ranging between 1 (for the plane circle) and n (see Section 4 below).

By taking account of the scale difference, we can exploit this similarity to estimate the writhing number without necessarily having to use model equations, but relying merely on curvature information; this may be useful for applications in laboratory experiments. To do this, one possibility is to define a scale factor h, given by

$$h = h_n(t) = \frac{[1 - [\cosh(nt)^2]^{-1}]}{n}, \tag{10}$$

and take $Wr(t) \approx \widetilde{W}r(t) = h_n(t)K(t)$. As we see from the bottom diagrams of Fig. 5, $\widetilde{W} = \widetilde{W}r(t)$ captures, on the whole, the essential features of the true writhe, with a localized jump in the absolute error $\Delta(Wr)$ due to the instantaneous, simultaneous formation of coils. Obviously, the choice of the scale factor h can be improved to have better control on the error.

Despite the approximation involved, this approach may turn out to be useful in many biophysical and chemical applications, where information on twist (rather than writhe) is quite relevant, this being related to the presence of particular molecular groups (hence to proteic and functional genomic properties) and to the energy contents of the elasto-mechanical and electro-chemical nature available. By approximating Wr with $\widetilde{W}r$, we can indeed estimate the twist from laboratory experiments according to the following working scheme:

(i) evaluate $K = K(t)$ from data analysis or direct inspection from experiments;

(ii) determine the scale factor $h = h_n(t)$ according to the value of n given by the total number of coils estimated in the molecule;
(iii) approximate $Wr(t)$ by $\widetilde{W}r(t) = h_n(t)K(t)$;
(iv) estimate Lk (for example, from information on topoisomerase reactions)—since Lk is a topological invariant, its value remains constant in time;
(v) evaluate $Tw(t)$ by $Tw(t) \approx Lk - \widetilde{W}r(t)$ (by Eq. (7)).

3. Energetics of folding

In terms of deformation energy, the process of folding and subsequent coiling of DNA filaments due to the relaxation of high twist is characterized by a transfer of torsional energy to bending energy [17]. The energetics of folding can be analyzed by using the model equations of Section 2. The DNA filament is modelled by a thin, inextensible rod, of uniform circular cross-section of area $A = \pi a^2$ ($a \ll L$) centred on $\mathcal{C}$. The elastic characteristics are specified by the bending rigidity K_b and the torsional rigidity K_t of the filament. In general, $\chi \in [1, 1.5]$, with $\chi = K_b/K_t$. In our case the stress–strain relation leads to the conventional quadratic form of the deformation energy, given by

$$E = E_b + E_t = \frac{1}{2}\oint_{\mathcal{C}} [K_b(c(\xi))^2 + K_t(\Omega(\xi))^2]\|\mathbf{X}'(\xi)\|\mathrm{d}\xi, \tag{11}$$

where we assumed zero natural twist rate of the filament fibers. The deformation energy E is thus given by the sum of the bending energy E_b, due to curvature effects, and the torsional energy E_t, due to torsion and intrinsic twist. If we also assume a *uniform* twist rate, i.e. $\Omega = \Omega_0$ constant, which is a reasonable assumption when we are in a state of minimum energy (away from inflexional deformations and energy localizations; see [12] for a detailed analysis of these particular situations), then, by (7) and (8), the torsional energy takes the form of the *mean twist* energy E_{tw}, that is

$$Tw = \frac{1}{2\pi}\Omega_0 L = \Omega_0, \quad \rightarrow \quad E_{tw} \equiv E_t|_{\Omega_0} = \frac{K_t L}{2}(\Omega_0)^2 = \pi K_t(Lk - Wr)^2, \tag{12}$$

which provides a useful relation for relaxed configurations. Note that the mean twist energy is proportional to the square of the linking number.

It is convenient to normalize everything with respect to a reference energy E_0, which we choose to be that given by the initial circular configuration of radius $R_0 = c_0^{-1} = 1$ and zero twist, that is

$$E_0 = \frac{K_b}{2}\oint_{\mathcal{C}} c_0^2 \mathrm{d}s = \pi K_b, \tag{13}$$

where s denotes here the usual arc-length. Thus, in the case of mean twist, the relative total energy is given by

$$\tilde{E}(t) = \tilde{E}_b(t) + \tilde{E}_{tw}(t), \tag{14}$$

where

$$\left.\begin{aligned} \tilde{E}_b(t) &= \frac{E_b(t)}{E_0} = \frac{1}{2\pi}\oint_{\mathcal{C}} (c(\xi,t))^2\|\mathbf{X}'(\xi)\|\mathrm{d}\xi, \\ \tilde{E}_{tw}(t) &= \frac{\pi K_t}{E_0}(Lk - Wr(t))^2 = \frac{(Lk - Wr(t))^2}{\chi}, \end{aligned}\right\}. \tag{15}$$

These two quantities, which together give information on the total deformation energy of the system, vary with time. One important question regards the existence of a 'deformation path' on which the total energy is decreasing monotonically. The existence of this path would favour folding, with the formation of coils. Since the total energy depends on the linking number Lk, which initially corresponds to pure twist (since at $t = 0$, $Wr = 0$), we expect to find a spectrum of critical twist values, given by Lk at $Tw = Tw_{\mathrm{cr}}$ and $Wr = 0$, that depends on the number N of coils formed, for which such a deformation path exists. Direct inspection of Eq. (14) reveals that this Lk is indeed strongly influenced by N, increasing with N. For a single coil, for instance, any $Lk \leq 8$ (Fig. 6(a)) gives a maximum

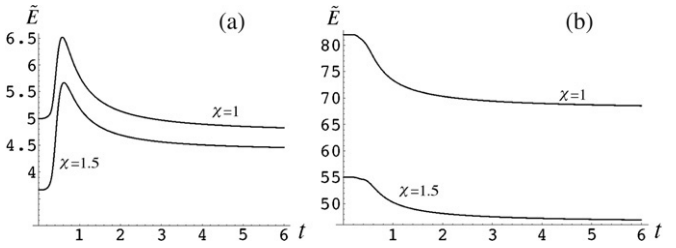

Fig. 6. Comparative behavior of the deformation energy (given by Eq. (14)) for the single coil case of Fig. 1, plotted against time for different values of Lk and χ: (a) $Lk = 2$, the maximum in total energy does not favour the formation of coiling; (b) $Lk = 9$, the monotonic decrease in total energy actually favours folding and filament coiling.

in the deformation energy that inhibits coiling. For $Lk \geq 9$, however, the total energy always decreases (Fig. 6(b)), independently of χ, thus favouring coiling. Note that, for given N, if the energy function is decreasing monotonically, then the largest value of energy is that attained when $t = 0$ (since energy is positive definite): hence, from Eq. (14), we have

$$\tilde{E}(0) = 1 + \frac{Lk^2}{\chi}, \tag{16}$$

giving information on the energy content absorbed by coiling. Simple estimates based on these diagrams can also give estimates on the energy share *per coil* produced, which appears to decrease with the number N of coils formed.

4. Hierarchical coiling and packing rate

Coiling provides an efficient way to store a large amount of filamentary information into tiny regions of space. Human DNA in chromosomes is compacted by a factor of about 10^4, and, when confined into a scroll of protein, has packing $D/L = O(10^{-5})$, where D and L denote typical sizes of the proteic region and DNA length [18]. A rough estimate of the maximum number of coils present in human DNA can be made by using helical geometry: from the definition of elementary arc of length δl of a helix of cylindrical radius r_0 and pitch p, we have

$$\delta l = \delta\alpha(r_0^2 + p^2)^{1/2}, \quad \text{and} \quad p_{\min} = a/\pi, \tag{17}$$

where $\delta\alpha$ is an elementary angle and a is the radius of the filament cross-section. Hence, from these relations the minimum length of a single coil is given by

$$l_{\min} = 2\pi\sqrt{r_0^2 + a^2/\pi^2}. \tag{18}$$

For an N-coil helical filament of length L, we must have $L = (N + 1)l$; hence

$$N_{\max} = \frac{L - l_{\min}}{l_{\min}} = \frac{L - 2\pi\sqrt{r_0^2 + a^2/\pi^2}}{2\pi\sqrt{r_0^2 + a^2/\pi^2}}. \tag{19}$$

According to typical values for human DNA [18], by taking $L = 10^{-2}$ m, $r_0 = 2 \times 10^{-8}$ m and $a = 5 \times 10^{-10}$ m, we have $N_{\max} = O(10^3)$–$O(10^4)$. For such a large N it is interesting to estimate the packing rates associated with higher-order helical coiling. Let us consider first the case of a single coil, as in Fig. 1: since L is kept constant, the average radius of the fully developed coiled curve will, in the limit $t \to \infty$, be half of the original. From the model equations, if N is the total number of coils produced, then we have:

$$L = 2\pi R = (N + 1)2\pi\rho \quad \to \quad \rho = \frac{R}{N + 1}, \tag{20}$$

where ρ denotes the average radius of curvature of the coiled state. If $N = N(t)$ is the number of coils produced per unit time, the *packing rate* $\rho = \rho(t)$ will be given by $\rho(t) = R/[N(t) + 1]$. This relationship is easily extended to a hierarchy of helical packings that are formed, starting from a fundamental structure of length, say, l_0 and the number

of coils N_0, to the final structure of total length L. The fundamental structure may be thought of as a primary helical structure wound on a secondary helical axis to form a secondary structure. This process can be extended to tertiary and higher-order structures, up to the resulting final structure. By taking account of these individual sub-structures, each contributing to the hierarchical helical coiling with length l_i and number of coils N_i ($i = 1, 2, \ldots$), we have

$$L = (N_1 + 1)l_1 = (N_1 + 1)(N_2 + 1)l_2 = (N_1 + 1)(N_2 + 1)\cdots(N_0 + 1)l_0. \tag{21}$$

An estimate based on kth-order coiling gives

$$L = O(\bar{N}^k l_0), \qquad \bar{N} = \Pi_k N_k; \tag{22}$$

hence, in general,

$$\rho(t) = O(R/N^k(t)), \tag{23}$$

which clearly shows a nonlinear dependence.

In conclusion, by analyzing a simple model of multiple folding and coiling, we have shown that a high degree of coiling does not necessarily involve high twisting or linking of the filament strands. Contrary to common belief and intuition, we have demonstrated that, by a suitable, but natural, choice of alternating strand passages, we can get as many coils as we want, while keeping the writhing number low. This efficient coiling process is perfectly admissible by natural mechanisms of folding, and it is energetically convenient. For sufficiently high initial twist, we showed that coiling is actually favoured by elastic energy relaxation, which is a result based on the deformation energy associated with the curvature and mean twist of the filament model. This aspect is related to the writhing instability of thin elastic filaments [19] and is responsible for the relaxation of highly twisted configurations to multiply folded states [17,20]. In the case of DNA molecules, where conformational information (due to geometry) and transcriptional mechanisms (associated with site-specific interaction of the filament strands) are inherently associated with local and global geometric properties, high values of writhe and twist, concentrated in short segments, as well as packing, surely play an important role in many functional and biological aspects. The results presented here, and the considerations above, are likely to be relevant in those contexts where there is a strong connection between efficient compaction, localization of energy, and functional processes.

Acknowledgements

R.L. Ricca acknowledges financial support from Italy's MIUR (D.M. 26.01.01, n. 13 "Incentivazione alla mobilità di studiosi stranieri e italiani residenti all'estero" and the COFIN2004–PRIN Project "Mathematical Models for DNA Dynamics $M^2 \times D^2$") and The Royal Society of London (Joint Project Grant under the European Science Exchange Programme).

References

[1] D. Baker, A surprising simplicity to protein folding, Nature 405 (2000) 39–42.
[2] M. Socolich, S.W. Lockless, W.P. Russ, H. Lee, K.H. Gardner, R. Ranganathan, Evolutionary information for specifying a protein fold, Nature 437 (2005) 512–518.
[3] B. Coulombe, Z.F. Burton, DNA bending and wrapping around RNA polymerase: A "revolutionary" model describing transcriptional mechanisms, Microb. Mol. Biol. Rev. 63 (1999) 457–478.
[4] P. Zhang, I. Tobias, W.K. Olson, Computer simulation of protein-induced structural changes in closed circular DNA, J. Mol. Biol. 242 (1994) 271–290.
[5] C. Behnam, D. Miller, Writhing and linking densities for closed, circular DNA, J. Knot Theory Ramifications 9 (2000) 577–585.
[6] F.B. Dean, A. Stasiak, T. Koller, N.R. Cozzarelli, Duplex DNA knots produced by *Escherichia coli* Topoisomerase, I. J. Biol. Chem. 260 (1985) 4975–4983.
[7] S.A. Wasserman, N.R. Cozzarelli, Supercoiled DNA-directed knotting by T4 topoisomerase, J. Biol. Chem. 266 (1991) 20567–20573.
[8] A. Stasiak, Circular DNA, in: J.A. Semlyen (Ed.), Large Ring Molecules, John Wiley & Sons Ltd., 1996, pp. 43–97.
[9] N.R. Cozzarelli, J.C. Wang (Eds.), DNA Topology and its Biological Effects, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1990.
[10] R.L. Ricca, Structural complexity, in: A. Scott (Ed.), Encyclopedia of Nonlinear Science, Routledge, New York, London, 2005, pp. 885–887.
[11] A.V. Vologodskii, N.J. Crisona, B. Laurie, P. Pieranski, V. Katritch, J. Dubochet, A. Stasiak, Sedimentation and electrophoretic migration of DNA knots and catenanes, J. Mol. Biol. 278 (1998) 1–3.

[12] F. Maggioni, R.L. Ricca, Writhing and coiling of closed filaments, Proc. R. Soc. A 462 (2006) 3151–3166.
[13] F.B. Fuller, The writhing number of a space curve, Proc. Natl. Acad. Sci. USA 68 (1971) 815–819.
[14] H.K. Moffatt, R.L. Ricca, Helicity and the Călugăreanu invariant, Proc. R. Soc. A 439 (1992) 411–429.
[15] G. Călugăreanu, Sur les classes d'isotopie des nœuds tridimensionnels et leurs invariants, Czechoslovak Math. J. 11 (1961) 588–625.
[16] J.H. White, Self-linking and the Gauss integral in higher dimensions, Amer. J. Math. 91 (1969) 693–728.
[17] R.L. Ricca, The energy spectrum of a twisted flexible string under elastic relaxation, J. Phys. A 28 (1995) 2335–2352.
[18] C.R. Calladine, H.R. Drew, Understanding DNA, Academic Press, 1992.
[19] E.E. Zajac, Stability of two planar loop elasticas, J. Appl. Mech. 29 (1962) 136–142.
[20] I. Klapper, Biological applications of the dynamics of twisted elastic rods, J. Comput. Phys. 125 (1996) 325–337.